NRL Report 8711



# A Technique of Comparative Analysis of Underwater Sound Transmission Loss Curves

G. R. GIELLIS AND B. B. ADAMS

*Acoustics Division*

June 28, 1983

NAVAL RESEARCH LABORATORY
Washington, D.C.

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|---|
| 1. REPORT NUMBER<br>NRL Report 8711 | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE *(and Subtitle)*<br>A TECHNIQUE OF COMPARATIVE ANALYSIS OF UNDERWATER SOUND TRANSMISSION LOSS CURVES | | 5. TYPE OF REPORT & PERIOD COVERED<br>Interim report on a continuing NRL problem<br>6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR*(s)*<br>G. R. Giellis and B. B. Adams | | 8. CONTRACT OR GRANT NUMBER*(s)* |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Naval Research Laboratory<br>Washington, DC 20375 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS<br>62759N<br>XF59552100<br>51-1517-0-3 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>Naval Electronic Systems Command<br>Washington, DC 20360 | | 12. REPORT DATE<br>June 28, 1983<br>13. NUMBER OF PAGES<br>27 |
| 14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)* | | 15. SECURITY CLASS. *(of this report)*<br>UNCLASSIFIED<br>15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Comparative statistical analysis
Sound transmission loss curves
Acoustic intensity curves
Intensity calculations

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

A frequent requirement in underwater acoustic experimentation and calculation is the comparison of detailed propagational loss curves. These curves may be characterized in term of three principal components:

- A monotonic trend, generally with an exponential decay
- A fluctuating component that contains a high percentage of periodic variation

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

20. ABSTRACT (Continued)

• A random component attributable to measurement error or unexplained environmental variation The quantitative comparisons are required to show equivalency of a theoretic and an experimental calculation, the degree to which alternative theoretic calculations agree, or equivalency of experimental replicates taken at different times. This report discusses this analysis in terms of the three components with the aid of comparative statistics. The trend analysis is accomplished with a linear regression on the logarithmically transformed range term. The residuals are examined with a turning point test and classed as systematic or random. The systematic residuals are examined with an autoregression procedure to estimate the systematic periodic component and to provide a period measure. A final turning point procedure is used to estimate the level at which systematic components have been found at a prescribed level of probability. The entire process provides a comparative analysis which is applied to several sets of experimental and theoretic acoustic data. The analyses show:

• The procedure is quite sensitive, allowing variations in the mean received level of 0.6 db to be readily distinguished. This enables the confident examination of records to ascertain correctness of total equipment calibration, flux density or calibration of sources or hydrophone or other matters that might produce an offset.

• Estimates of periodicity with high degrees of confidence were found capable of distinguishing variations of 2 in 60 km. These would indicate error in the mean velocity profile, or the theoretic method, either of which could alter the range scale of the estimated periodic propagation loss curve.

• Distinctions between types of models and the degrees of smoothing were quite readily and reliably made. This enables deficiencies in the matching environmental parameters, or the deficiencies in the calculational procedures to accommodate various factors, to be readily perceived. All of the procedures are based on published references and are available for immediate application in underwater acoustic cases as well as other areas which result in experimental and/or theoretic curves that can be decomposed into the three components of trend, periodic variation, and random residual.

## CONTENTS

# A TECHNIQUE OF COMPARATIVE ANALYSIS OF UNDERWATER SOUND TRANSMISSION LOSS CURVES

## INTRODUCTION

A number of computer programs for intensity calculations are now widely available to scientists engaged in underwater sound studies. These programs produce curves which indicate transmission loss as a function of range, comparable to the type derived from available experimental data. Some of the more sophisticated models, such as TRIMAIN used at NRL, can handle horizontal variations in sound speed and include bottom topography, and produce four different types of intensity calculations.

There is a need for a systematic comparison of true intensity curves predicted by a model with those obtained by field measurements. Such a procedure would facilitate the modeling process by providing quantitative measures of model adjustment effects. The objective of this study was to develop an analysis procedure capable of meeting this requirement.

Acoustic intensity curves have three basic components: a long-term trend, oscillations about this trend and residual random effects. One or more of these components may not be present to a significant degree, depending on a given physical situation. The procedure we have developed is designed to establish the existence of the components, and to isolate them for separate examination and quantitatively estimate their contribution. An outline of the recommended procedure follows. The procedure has been exercised on acoustic models and experimental data, with detailed results given in the second section of this report. The report ends with conclusions regarding the progress to date in the development of the procedure and recommendations for further study. The appendix of this report is devoted to a complete description of the procedure which includes specific formulas and a discussion of underlying assumptions.

## PROPOSED ANALYSIS PROCEDURE

Given an intensity curve $X(r)$, the long-term trend is assumed to be of the elementary form $X_L(r) = A + B \log r$ (Fig. 1a). The A and B coefficients are determined by least squares formulas. The residuals constitute a derived curve, $X'(r)$ (Fig. 1b). The subsequent tests used depend upon the presence of significant randomness in $X'(r)$, as measured by a turning point test [1]. If the residuals are random, the intensity curve, $X(r)$, is described by only two components; the long-term trend, $X_L(r)$, and the random residual, $X'(r)$. To compare two-component curves of this type, similarity tests based on confidence intervals for A and B and an estimate of the standard deviation for $X'(r)$ can be used. We can also compare curves by examining the distributions of variance between the $X_L(r)$ and $X'(r)$ components. If the curve $X'(r)$ fails the test for randomness, then a third significant, oscillatory, component exists. In this case, the subsequent comparisons of trend coefficients can be made disregarding the oscillatory component with a small loss in comparison precision. Alternately, if full compliance with statistical assumptions is necessary, the oscillatory component can be removed and a second regression made for refined trend parameter estimates.

In a large percentage of cases, transmission-loss curves possess a strong oscillatory component. We have assumed it to be of the form

$$X_0(r_k) = \sum_{j=1}^{n} a_j \, X'(r_{k-j}).$$

Manuscript approved February 24, 1983.

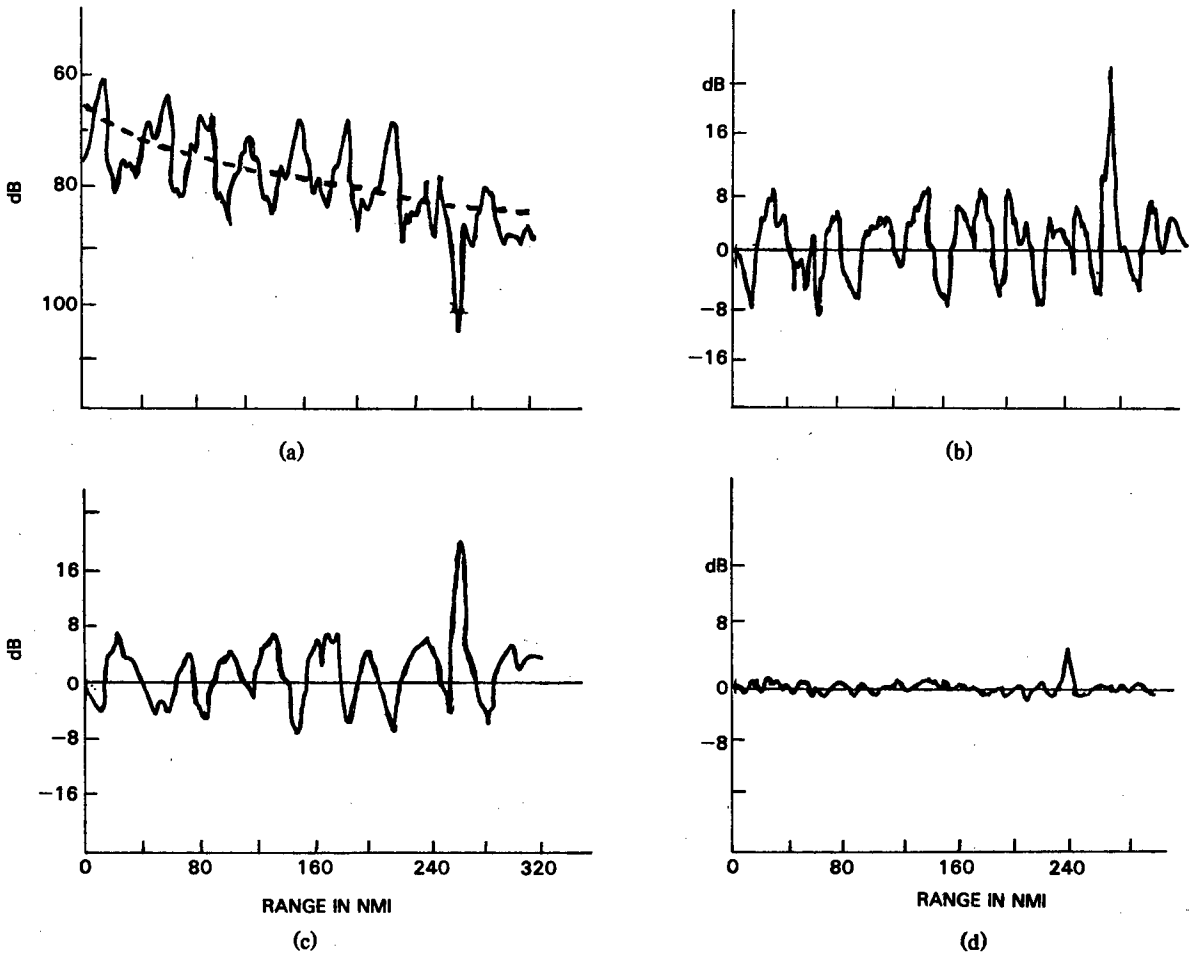



Fig. 1 — (a) A three-component propagation loss curve with an exponentially decaying average overlay (dashed line); (b) the derived oscillatory curve, $X'(r)$, after the trend in Fig. 1(a) is removed; Fig. (c) an autoregressive estimate, $X_0(r)$, of the residual oscillations in Fig. 1(b); Fig. (d) the final residual, $X_R(r)R$, equaling the difference between the oscillatory residual of Fig. 1(b) and the oscillatory estimate of Fig. 1(c).

Calculation of the coefficients $[a_j]$ is discussed in the appendix of this report. Briefly, it requires the solution of a system of equations involving the autocorrelation function for $X'(r)$. The autocorrelation function is also used to estimate the principal period of a transmission-loss curve, such as the convergence zone period, and further, to calculate a zone-spacing ratio, designed to compare the oscillation periods of two curves. To show whether the autoregressive scheme is complete, the residual component $X_R(r)$ is obtained as $X_R(r) = X'(r) - X_0(r)$ (Fig. 1c,d). At this stage, a turning point test is again applied to see if $X_R(r)$ satisfies a randomness criterion. If not, refinements are necessary in the autoregressive fit procedure.

After the separation into components has been accomplished, curves for model or experimental data can be compared for the distribution of variance among these components. The comparisons are quantitative, reproducible, and contain probability thresholds, or confidence intervals all of which can be used for systematic comparison of data sets, model sets or data/model tests.

## A COMPARATIVE ANALYSIS OF UNDERWATER SOUND TRANSMISSION LOSS CURVES

This section presents results obtained from use of the analysis procedure for a model/experimental comparison study. The objectives are to:

- Examine and show for example the practical application of the procedure.
- Test the resolution of the methods on two similar data sets arising from slightly different conditions.
- Test the resolution of the method in a comparison of data vice against model performance.
- Test the resolution on model curves to distinguish algorithm differences.

The experimental data were obtained during a controlled run of 300 nmi by the *USNS Mizar*, directed by excellent satellite navigation. The signal sources were small explosive charges carefully timed with synchronized WWV clocks for precise range and depth control. The shot spacing was 1/2 nm and the depth 91.4 m. Sound speed profiles were measured at the ends and in the track center; detailed bathymetry was measured throughout the run. The signals were received by hydrophones suspended from two ships, the R/V *Knorr* and *USNS Gibbs*, stationed at the beginning and end of the track, respectively. The *Knorr* phones were vertically separated by 150 m, with the lower unit at a depth of 3386 m. The same arrangement was used for the *Gibbs* phones, with the bottom unit at a depth of 3020 m. Transmission-loss curves were computed for two low-frequency third-octave bands, separated by 50 Hz. Table 1 shows the labeling systems used for the curves.

Table 1 — Experimental Intensity Curves

| Curve | Ship | Hydrophone | Frequency |
|---|---|---|---|
| KXUL | KNORR | Upper | Low |
| GXUH | GIBBS | Upper | High |

A series of intensity curves to be used for measured and predicted comparisons were generated largely by the computer program TRIMAIN using measured bathymetry and sound-speed profiles, along with the appropriate source and receiver depths. Using incoherent Type I, summation intensity calculations [1], three curves corresponding to the experimental runs were generated. A second, Type II method [2] using a ray weighting based upon an exponential probability distribution function in depth was used on two program runs. A third method used a Lloyd Mirror (LM) correction for proximity to the surface. A listing of the resultant TRIMAIN curves used for our analysis is given in Table 2.

Table 2 — TRIMAIN Model Intensity Curves

| Curve | Ship | Phone | Type | Frequency |
|---|---|---|---|---|
| KTUI | KNORR | Upper | I | |
| KTUL (LM) | KNORR | Upper | I | Low |
| GTUI | GIBBS | Upper | I | |
| GTUII | GIBBS | Upper | II | |

In addition, one run was made with the Fast Asymptotic Coherent Transmission model (FACT) to match the first 250 values of the curve KXUH. The FACT Program contains a first-order caustic computation but is restricted to a single sound-speed profile and flat horizontal bottom. We denote the intensity curve for this case as KFUH and, for comparison, use only the first 250 values of the corresponding experimental curve, denoted as KXUH(F).

In the remaining articles we discuss the results obtained as the analysis procedure was applied to the experimental and model curves listed above. We have selected two groups of curves: The first consists of KXUL, KTUI, and KTUL (LM) (Figs. 2a,b,c); that is, an experimental curve for the *Knorr* data, with two corresponding TRIMAIN runs, differing in the type of intensity calculations used. The second group is a similar selection, comprising the curves GXUH, GTUI, and GTUII (Figs. 3a,b,c), based on the *Gibbs* data.

## LONG-TERM TRENDS

Following the discussion of the appendix of this report, the long-term trend is assumed to be of the form A + B log ($r$), and least-squares Eqs. (A2) and (A3) are used to calculate the coefficients A and B. The residual curves remaining after trend removal are denoted with a prime (') superscript. Thus,

$$KTUI' = KTUI\ (r_k) - A - B \log\ (r_k). \tag{1}$$

The long-term trend and the standard deviation of the residual curve are used to define the data trend bands shown in Figs. 2 and 3. Note that each figure displays an experimental curve and two model curves. The bands are defined by displacing the long-term trend of the experimental curve upward and downward. The distance moved is two standard deviations, as calculated from the residual of the experimental curve after long-term trend removal.

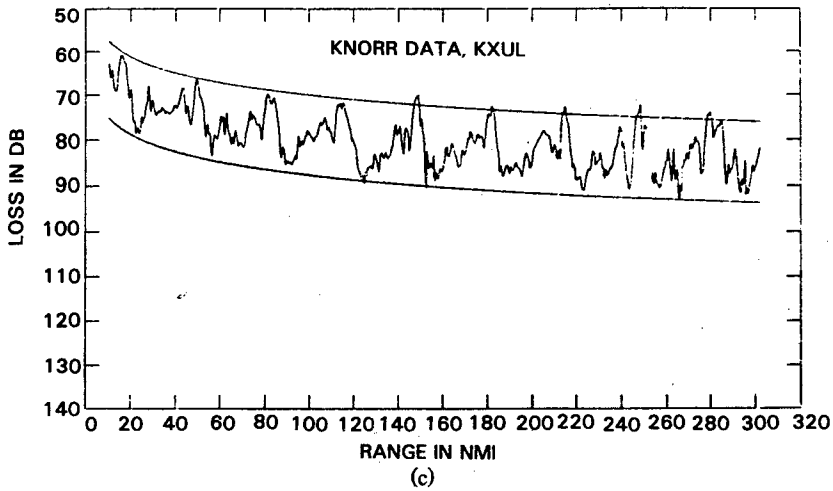
A band-fit coefficient can be calculated as a measure of the goodness-of-fit of the model curves. It consists of the percentage of band-confined model data points divided by the percentage of experimentally measured data points falling within the band.

Table 3 tabulates the data trend bandwidth and band-fit coefficients to provide a simple numeric comparison of model-generated and field-measured data. This test is not of the same rigor as the statistical measurements tabulated in table 4 below. However, it is recommended for display and presentation in conjunction with the type of plots shown in Figs. 2 and 3.

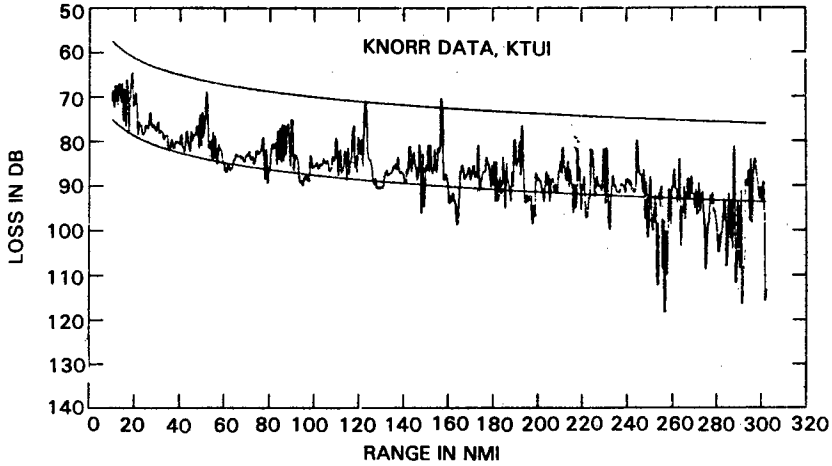
Table 4 lists the results of the trend estimation on several model and two experimental data suites. The significant features of this compilation are:

- The mean values of the data sets are different and ordinarily would reflect systematic bias in a entire suite under comparison, a calibration error or possibly a bottom condition at a near bottomed receiver not adequately modeled. The confidence interval for this mean is included, for statistical comparison, in this grouping.

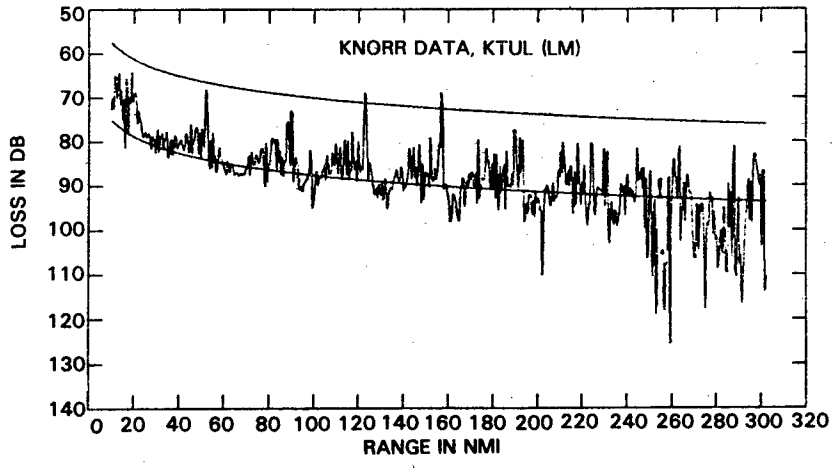
- The regression coefficient, B, shows the estimated exponential power decay of the sets. In the first *Knorr* group, we see a distinctly sharper fall (larger exponent) in the two model sets. These model runs were included to show how a modeling error, ray drop out, purposely produced and plotted Fig. 2b,c, can produce a definite measurable difference. The effect is also reflected in the mean value difference. Next, considering the *Gibbs* suite, we see a case of strong smoothing (GYTUII) suppressing the growth of the decay constant, B, and also increasing the model mean to near the observed set seen in GXUH. While the range smoothing has been deliberately overdone for illustration, it is clear that models could be brought into correspondence by this method with data, and more discriminating tests for spectral content might be required for distinct numeric separation.




(c)

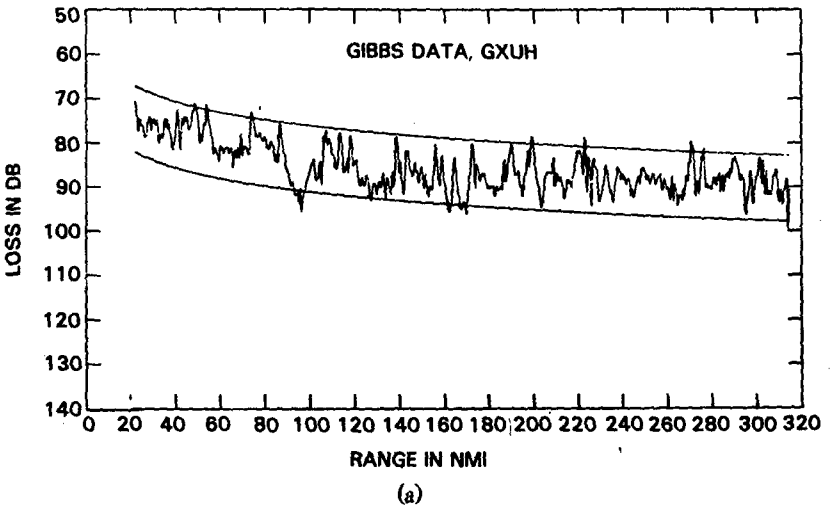



(b)

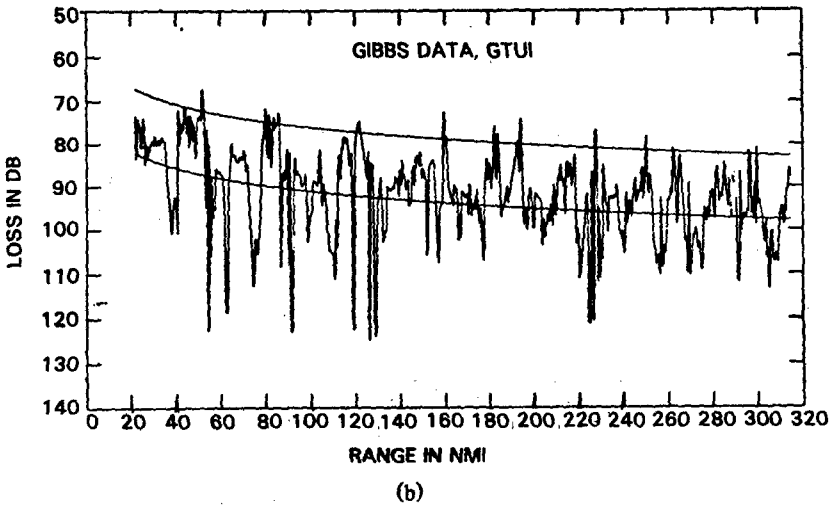



(a)

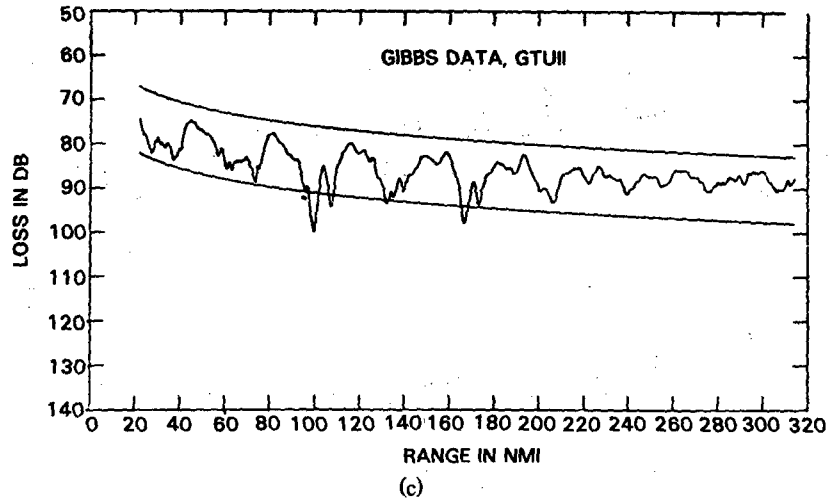
Fig. 2 — (a) The measured and fully corrected estimates of propagation loss taken on board the R/V KNORR; (b) TRIMAIN model calculations employing the bathymetric and velocity profile information acquired during the measurement of Fig. 2(a); (c) A second TRIMAIN calculation followed by the application of a Lloyd Mirror correction; (both model calculations are produced with deliberate flaws in the number of rays included beyond 250 km to provide material for the analysis technique to discriminate).




(a)




(b)




(c)

Fig. 3 — (a) The measured and fully corrected estimates of propagation loss taken on board the USNS GIBBS; (b) model calculations utilizing TRIMAIN and employing the same bathymetric and velocity profile information acquired during the measurement of Fig. 3(a); (c) A second TRIMAIN calculation followed by the application of a Lloyd Mirror correction.

Table 3 — Data Trend Bandwidth

| Exp. Curve | Model Curve | Data Trend Bandwidth (dB) | Band Fit. Coeff. |
|---|---|---|---|
| KXUL | KTUI | 17.6 | .78 |
| KXUL | KTUL(LM) | 17.6 | .62 |
| GXUH | GTUI | 15.0 | .64 |
| GXUH | GTUII | 15.0 | 1.01 |
| KXUH(F) | KFUH | 16.5 | 1.00 |
| KXUH(F) | KTUI(F) | 16.5 | 1.03 |

Table 4 — Results of Long-term Trend Removal

| Curve Designate | Mean M | Conf. 95% M± | Slope B | Conf. 95% B± | Std. Deviation | |
|---|---|---|---|---|---|---|
| | | | | | Data,S | Resid. $S_P$ |
| KXUL | 90.1 | .5 | 12.6 | 1.1 | 6.1 | 4.4 |
| KTUI | 96.8 | .6 | 17.6 | 1.3 | 7.8 | 5.2 |
| KTUL(LM) | 98.1 | .7 | 18.5 | 1.5 | 8.7 | 6.1 |
| GXUH | 95.6 | .4 | 13.6 | 1.1 | 5.4 | 3.8 |
| GTUI | 101.5 | .8 | 14.8 | 2.7 | 10.2 | 9.3 |
| GTUII | 95.9 | .4 | 9.5 | 1.0 | 4.3 | 3.4 |
| KXUH(F) | 88.2 | .8 | 16.3 | 1.8 | 6.2 | 4.1 |
| KFUH | 88.9 | .8 | 19.2 | 1.6 | 6.5 | 3.6 |
| KTUI(F) | 91.3 | .7 | 15.1 | 1.6 | 5.6 | 3.6 |

• In the third set in Table 4, we have a comparison of the FACT model with data. A slight bottom-loss adjustment would probably raise the mean and decrease the decay constant, B, to near perfect coincidence. One advantage this last set shows in model/data comparisons is how range constraint improves the quality of the match. The last curve, KTUI(F), is a TRIMAIN estimate run to the same 250-range-point limit of the FACT model and quantitatively shows that at lesser ranges, the ray density is adequate, and the model improves. Generally, as might be expected, long-range predictions and comparisons prove the most difficult and are likely to require the techniques described in this report.

• The last two columns of Table 4 show the original variance and the remaining or residual variance. This last column, in particular, illustrates the effects of smoothing in the GTUI/GTUII contrast. A variance comparison test, such as the F test discussed in Ref. 2, is ideal for quantitative smoothing comparisons or processing bandwidth comparisons.

Following a set of qualitative comparisons as previously described, assume that we have further noted and examined for cause the difference in mean and coefficient estimates and noted the confidenced intervals on each. More detailed comparisons of two data suites require the following additional calculations:

- Generally for curve parameter comparison, it is essential for the data to originate from the same population. This can be tested by forming the F ratio of the residual variances of each curve pair. Approximate similarity is usually sufficient.

- Using a pooled residual variance, a standard deviation of the difference of the regression coefficient, (decay constant) is computed.

- A confidence interval in this difference variance is then computed using the T distribution.

While each of these steps is described in standard texts on statistics, a factor not immediately apparent is that these three steps can be approximated accurately as follows:

- If the variances are nearly the same, assume the populations are the same.

- Usually most experimental model comparisons will involve large numbers (50 or more) points, and pooling for more accurate variance estimation is marginally useful and may be ignored.

- The confidence coefficient for the difference in two coefficients is simply computed as the square root of the sum of the squares of the two subject coefficients.

As an example of the preceding procedure, Table 4 shows Gibbs data, GXUH, whose residual variance is 3.8 dB. The TRIMAIN model with range averaging, GTUII, gives 3.4 dB. Let us assume these are equal. The 95% confidence interval halfwidth for B is about 1.1 in each case which gives combined (root of the sum of the squares) difference halfwidth of 1.6. The difference in the coefficients, however, is 4.1, that is, 13.6 - 9.5. This greatly exceeds our 95% interval and we may say the probability is less than one in twenty that the curves are the same. In this instance, the model parameters definitely need adjustment.

The simplified technique also can be used to compare the means of two groups. Using the same *Gibbs* data/model, GXUH/GTUII, Table 4, we have for the 95% confidence interval on the difference in means, a root-of-sums squared of 0.6 which is not exceeded by the 0.3-dB difference in means. Thus, the smoothing brought the mean under control, reduced both the data variance (10.2 to 4.3), and the residual variance (9.3 to 3.4), a small amount more than required, but rendered the slope unsuitable, (14.8 to 9.5).

A conclusion of this comparison is: less smoothing and some physical factor related to mean offset probably require consideration.

The second phase of the trend analysis procedure requires the residual curves, $X'$ to be tested for randomness with the turning point test (see Appendix A.2). Based on the hypothesis that the curve is random, confidence limits for the number of turning points are calculated, using Eq. (A.8). Table 6 gives these limits, and the observed count of turning points for the six selected curves whose plots will be examined.

In each case, the number of turning points falls outside the limits. Thus, we reject the hypothesis of randomness and conclude that each of the curves in Table 6 has a significant oscillatory component.

## OSCILLATORY RESIDUAL CURVE ANALYSIS

The turning point test for randomness establishes the existence of significant oscillations in the trend residual. Each of the model and experimental residual curves which are given in Figs. 4-7, exhibit this strong oscillatory component. To begin the analysis, we may express one of the residual curves, $X'(r)$ as

$$X'(r) = X_0\,(r) + X_R\,(r). \tag{2}$$

Here $X_0(r)$ is the oscillatory component, and takes the form of an autoregressive scheme,

$$X_0(r_k) = a_1 X'(r_k - 1) + a_2 X'(r_k - 2) + \ldots + a_m X'(r_{k-m}). \quad (3)$$

The final curve, $X_R(r)$, will then be a purely random sequence.

The procedure for autoregressive scheme fitting, as discussed in Ref. 3, involves the choice of an order $m$, and the calculation of the coefficients $a_1, a_2, \ldots, a_m$ as the solution of a system of equations defined by the autocovariance function of the curve $X'(r)$. A FORTRAN computer code of the type devised by Robinson [4 Section 2.8] was used for this purpose. To provide a measure of completeness, the normalized mean square error, $E_m$, is calculated as the ratio of the residual variance to the trend variance, the program estimates $E_k$ for all orders $k$ less than or equal to $m$ and calculates the coefficients $a_1, a_2, \ldots a_m$. To provide an accurate estimate of the residual variance for a variety of curve types, a large value for $m$ is recommended. After several trials, the value $m = 128$ was sufficient for essentially all cases while requiring less than two minutes of machine time. In running the program for this order, the differences between values of $E'_m$ and $E_{m-1}$ were less than 0.002 in all cases, indicating that the residual variance had reached a stable level. After the autoregressive scheme fit has been made, the final residual component is tested for randomness by using Kendall's turning point test.

Table 5 lists the essential information obtained in the autoregressive analysis. The first column gives the normalized mean square error at $m = 128$. This is followed by the 95 percent confidence intervals for the turning point test, along with the observed number of turning points for each residual curve. In each case, this value falls within the confidence limits, and we can accept the hypothesis of a random residual curve. Figures 6 and 7 show the autoregressive scheme fits which were calculated as the oscillatory component of six representative curves.

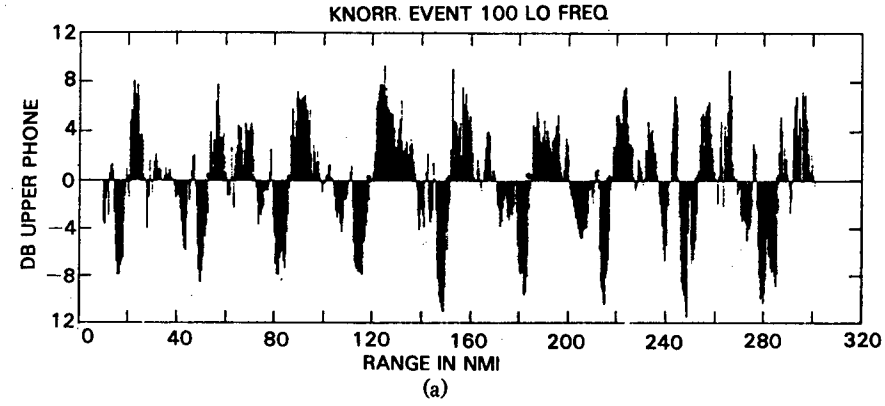
Table 5 — Separation of Zero-Mean Curves into Oscillatory and Residual Components

| Curve | Mean Sq. Error $E_{128}$ | Confidence Intervals for Turning Points | | Observed No. of Turning Points |
|---|---|---|---|---|
| KXUL' | .119 | 361 | 399 | 366 |
| KTUI' | .606 | 361 | 403 | 402 |
| KTUL(LM)' | .623 | 365 | 403 | 387 |
| GXUH' | .337 | 359 | 397 | 364 |
| GTUI' | .518 | 365 | 403 | 371 |
| GTUII' | .018 | 336 | 373 | 344 |
| | $E_{64}$ | | | |
| KXUH(F) | .153 | 150 | 174 | 159 |
| KFUH | .093 | 139 | 162 | 140 |

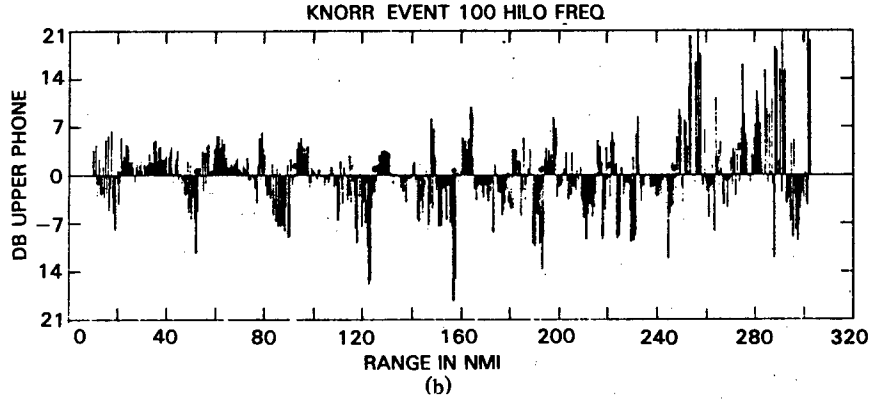
The autocorrelation function calculated as part of the preceding above procedure can be used to estimate the principal period of the oscillatory component. This interval is calculated from tabulation of the correlation function and is the interval between successive maxima. Usually several cycle peaks will be clearly evident and the average computed will accurately reflect the chief periodic phenomenon. In the case of all the present examples, this is the convergence zone cycle. This cycle can be used to define the zone cycle ratio, $ZC = (C_1 - C_2)/C_1$ to provide fractional error comparison of the curves with periodicities. Table 7 shows cycles in nautical miles and the period ratios for the several model runs as compared with the two sets of experimental data.

Table 6 — Test for Randomness at 95 Percent Confidence Interval

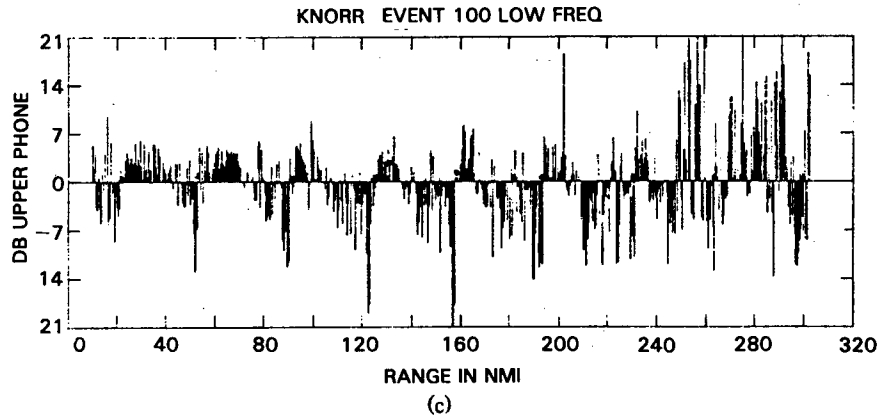
| Curve | Confidence Limits | | Number of Turning Points |
|---|---|---|---|
| | Lower | Upper | |
| KXUL′ | 362 | 401 | 214 |
| KTUI′ | 360 | 400 | 357 |
| KTUL(LM)′ | 364 | 404 | 349 |
| GXUH′ | 345 | 383 | 272 |
| GTUI′ | 363 | 402 | 309 |
| GTUII′ | 337 | 374 | 109 |




(a)

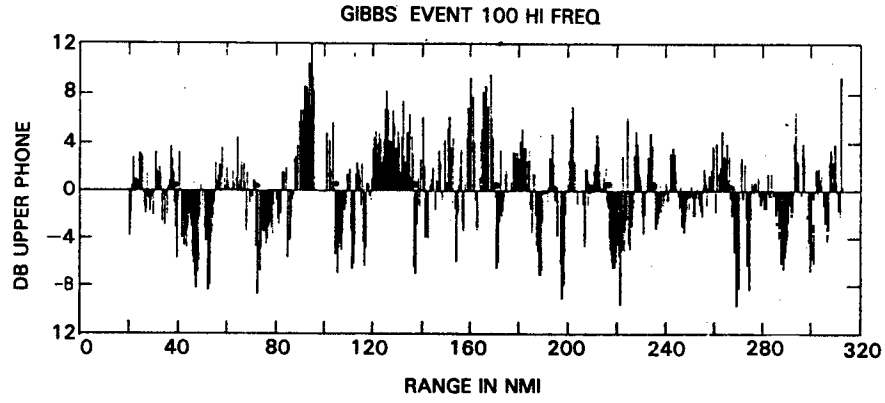



(b)

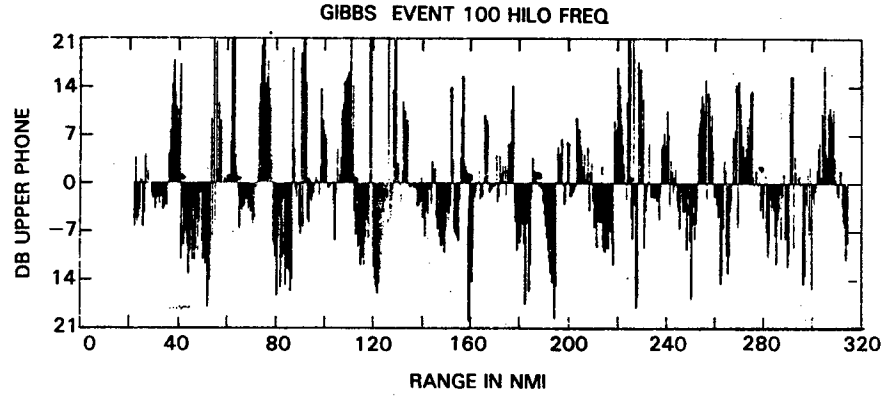



(c)

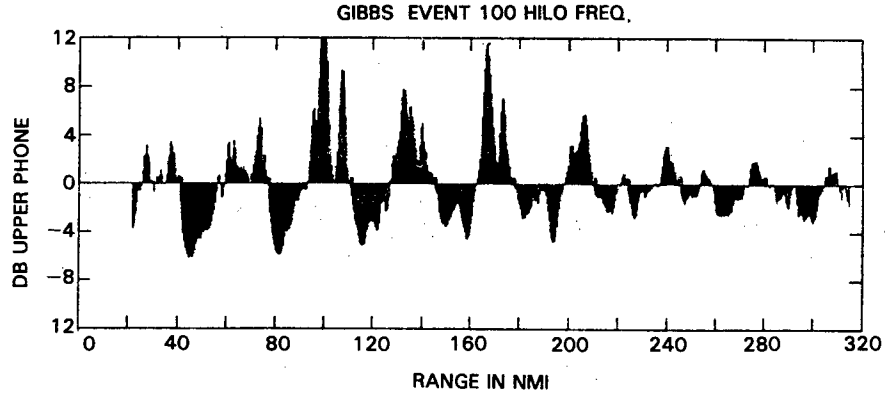
Fig. 4 — Three derived residual curves, $X'(r)$, from the respective R/V KNORR data and model calculations illustrated in Fig. 2.




(a)

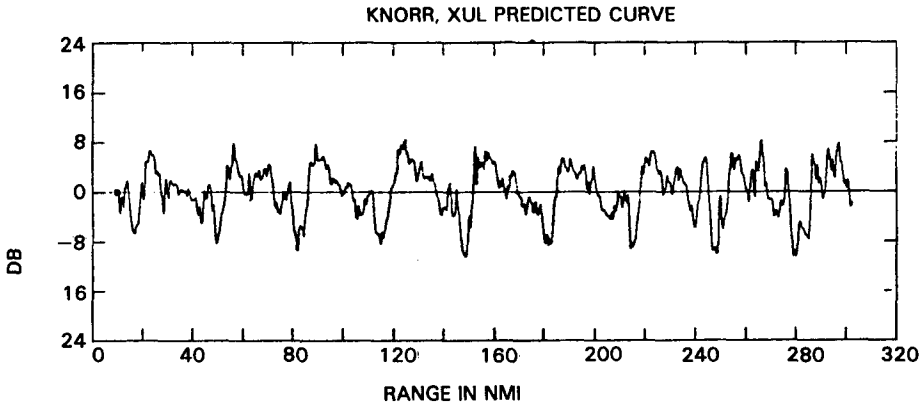



(b)

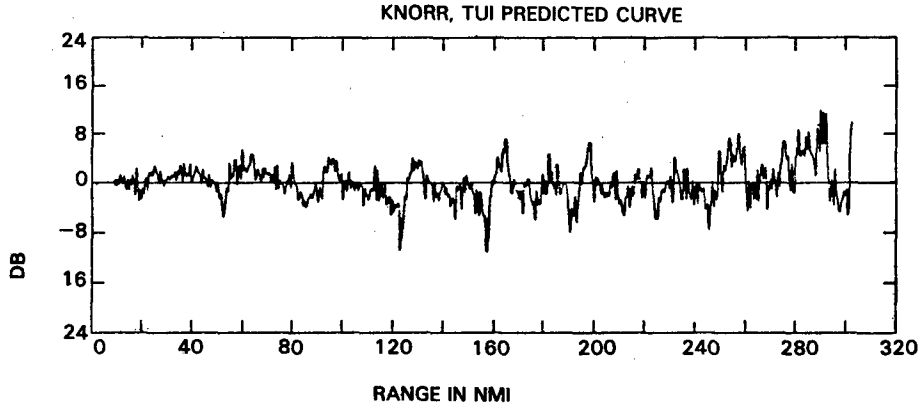



(c)

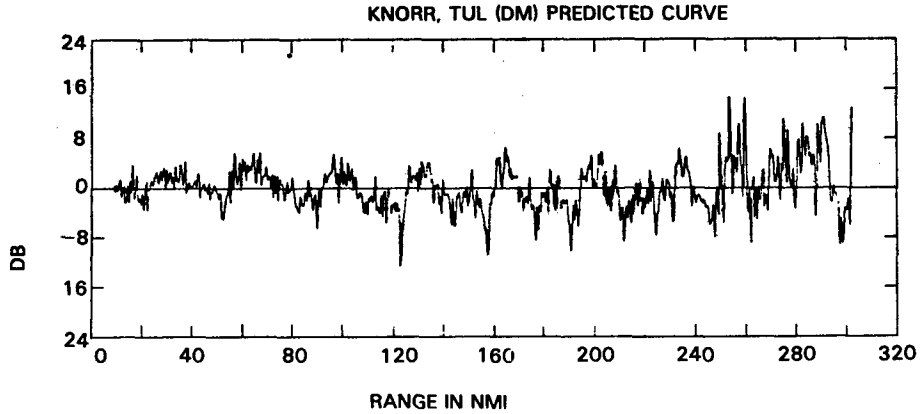
Fig. 5 — Three derived residual curves, $X'(r)$, from the respective USNS GIBBS data and mode calculations illustrated in Fig. 3.




(a)

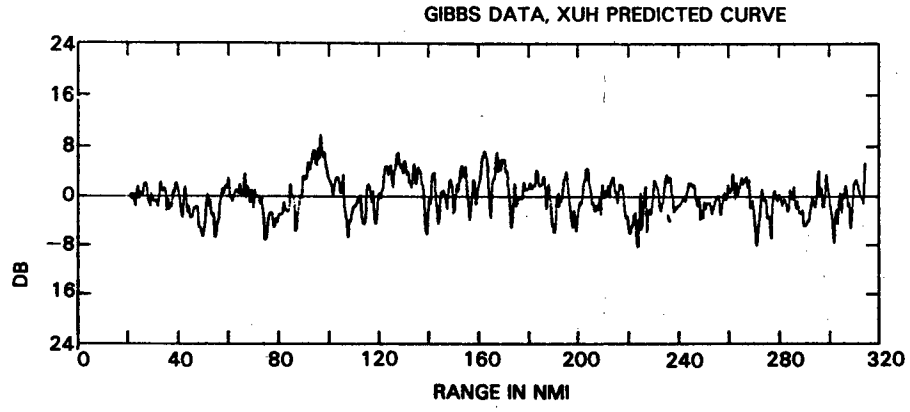



(b)

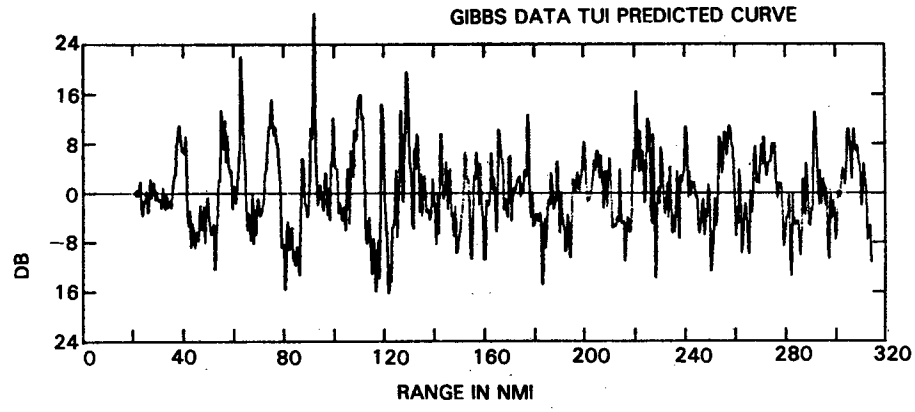



(c)

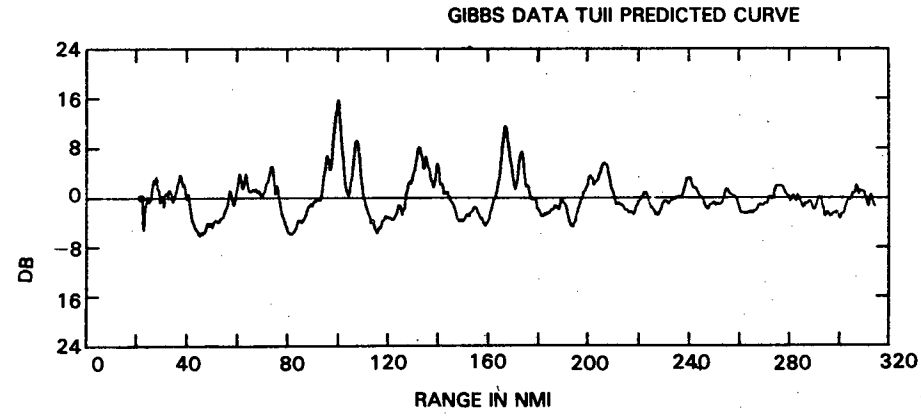
Fig. 6 — Three autoregressive estimates of the residual components respectively illustrating R/V KNORR data and model calculations shown in Fig. 2 and 4.




(a)




(b)




(c)

Fig. 7 — Three autoregressive estimates of the residual components respectively illustrating USNS GIBBS data and model calculations shown in Figs. 3 and 5.

Table 7 — Zone Cycles and Zone Spacing Ratios for Oscillatory Components

| Exp. | Curves $C_1$ (nmi) | Curves Model | Curves $C_2$ (nm) | $Z_R$ |
|---|---|---|---|---|
| KXUL | 33.1 | KTUI | 35.1 | −0.060 |
| KXUL | 33.1 | KTUL(LM) | 34.1 | −0.030 |
| GXUH | 32.6 | GTUI | 34.1 | −0.076 |
| GXUH | 32.6 | GTUII | 34.6 | −0.061 |
| KXUH(F) | 31.3 | KFUH | 36.0 | −0.150 |

The values shown in Table 7 indicate that the sample model cycles are greater than the experimental. The Lloyd's mirror calculations and Type II representations do not change the principal cyclic structure significantly as this is a fundamental characteristic of each measured or predicted curve which is not affected even by a heavy smoothing operation.

**Comparison of Variances**

At each stage of the separation process, estimates were made for the variance of the component curves, using the familiar formula,

$$V = \frac{1}{N-1} \sum_{k=1}^{N} (X(r_k) - \bar{X})^2. \tag{4}$$

Here $\bar{X}$ is the mean of the curve, and $N$ is the number of range values. In this manner, we arrived at estimates of the variances, $V$ for the initial curve, $VL$ for the long-term trend, $V'$ for the trend residual curve, $V0$ for the oscillatory component, and $VR$ for the final residual. Table 8 lists these values for our illustrative set of curves.

Table 8 — Variances for Component Curves

| Curve | $V$ | $V_T(= V_L + V_0 + V_R)$ | $V_L$ | $V'$ | $V_0 + V_R$ | $V_0$ | $V_R$ |
|---|---|---|---|---|---|---|---|
| KXUL | 37.30 | 37.05 | 17.99 | 19.31 | 19.06 | 16.86 | 2.20 |
| KTUL(LM) | 75.43 | 72.49 | 38.44 | 36.99 | 34.05 | 13.28 | 20.77 |
| KTUI | 61.53 | 59.31 | 34.76 | 26.76 | 24.55 | 9.41 | 15.14 |
| GXUH | 29.13 | 28.86 | 15.10 | 14.03 | 13.76 | 9.14 | 4.62 |
| GTUI | 103.36 | 101.78 | 17.75 | 85.61 | 84.03 | 40.39 | 43.64 |
| GTUII | 18.75 | 18.72 | 7.40 | 11.35 | 11.32 | 11.10 | 0.22 |
| KXUH(F) | 38.34 | 38.02 | 21.25 | 17.09 | 16.77 | 14.28 | 2.50 |
| KFUH | 41.56 | 42.35 | 29.54 | 13.01 | 12.81 | 11.56 | 1.25 |

Thus, if we have separated the initial curves into independent components, we should have $V' = V0 + VR$, and $V = VL + V0 + VR$. In practice, the results were very close to the theoretical, with the largest discrepancy about 8 percent of the amount involved. Table 9 lists the proportions of the initial variance which can be attributed to each of the three components with the fractions normalized to the total for each curve.

Table 9 — Distribution of Variances

| Curve | $P_L = V_L/V_{TOT}$ | $P_0 = V_0/V_{TOT}$ | $P_R = V_R/V_{TOT}$ |
|---|---|---|---|
| KXUL | .486 | .455 | .059 |
| KTUL(LM) | .530 | .183 | .287 |
| KTUI | .586 | .159 | .255 |
| GXUH | .523 | .317 | .160 |
| GTUI | .174 | .397 | .427 |
| GTUII | .395 | .593 | .012 |
| KXUH(F) | .559 | .376 | .066 |
| KFUH | .696 | .273 | .030 |

In summarizing the observed three part variance distribution of our sample set, a number of observations can be made:

- Both of the experimental data sets show a comparatively small amount of residual variance that is only approached by the FACT model operating on a restricted range of data and the smoothed TRIMAIN model on the whole range. In both model instances, extreme excursions are controlled; this parallels the frequency domain averaging that is a feature of typical (1/3) octave propagation data acquired with explosive charges. We would expect a measurement made with a narrow band source to more closely match the random variability of model data (excluding the ray dropout cases included here only as negative examples).

- A strong periodic component, developed from convergence zones in the present data, and not unusual in other instances, can be expected to be always present. Simple bottom-limited propagation would be a common case not likely to show periodic components.

- An elementary point is that the variance distribution such as observed in Table 9 is clearly affected by the proximity of the first point to the origin; starting at greater ranges, the first term would be smaller in all cases.

The statement of the guiding nature of these quantitative measures must be reiterated. The three summary remarks show how each of the components as well as their distribution are affected by measurement techniques, range, and computational procedures. Strong conclusions can be drawn in specific cases and these can be strongly supported and have considerable sensitivity; however, the methods are not automatic and their application is supportive to the analyst who retains responsibility for their correct application and results.

## CONCLUSIONS AND RECOMMENDATIONS

A sequence of known statistical procedures has been gathered and used on the comparative analysis of measured and calculated propagation loss data. The procedures deliberately have been kept simple to promote widespread use with a minimum of computer or calculator expense. The cases chosen for examples in Table 3, show the 95% confidence interval on the mean of the sets is on the order of 0.6 dB even though a number of the model runs were purposefully flawed for illustration. This sensitivity for calibration checks; flux density estimates and hydrophone calibrations, etc. were unexpected. Similarly the exponential decay constant 95% confidence intervals, or slopes, were of order 0.15 where 20 would be spherical spreading. All the cases were readily distinguishable. The distribution of variance in the tested cases shown in Table 9 also showed marked distinctions between

model types as well as experimental data. The convergence zones of the chosen sample data were remarkably periodic so that in the two examples the oscillatory and long-term trend were near equal in power, and the final random residual variance was only 6 and 16% in the two cases. The model results were deliberately not tuned to the experimental data to better reflect what a first application of the methods would produce. As a consequence, most of the model outputs contained a much larger random residual component which was suppressed only in the smoothed cases. This is similarly, in retrospect, not unexpected since the computer models are comparable to continuous wave (very narrow band) data and the experimental results have one third octave frequency domain averaging. In Table 7 shows another result where comparisons of measured and model convergence cycle lengths are listed. All the model cycle lengths exceed all the experimental lengths. The discrepancies are small, 3%, but consistent and estimated to result from velocity profile error.

The major objective has been to illustrate the surprising power of a sequence of comparatively elementary procedures and the major recommendation is to use objective measures such as those discussed in general experimental and analytic studies.

## REFERENCES

1. M.G. Kendall, *Time Series*, Hafner Press, New York, 1973.

2. E.J. Hanan, *Time Series Analysis*, Methuen, London, 1960.

3. C. Chatfield and M.P.G. Pepper, "Time-Series Analysis: An Example from Geophysical Data", Royal Stat. Soc. J (C), 20, 217-238 (1971).

4. G.M. Jenkins and D.G. Watts, *Spectral Analysis and Its Applications*, Holden-Day, 1968.

# Appendix
# THE ANALYSIS PROCEDURE—A DETAILED DESCRIPTION

## COMPUTATION OF THE LONG-TERM TREND

The threefold separation of major contributing components of a transmission-loss curve initially uses the form $XL(r) = A + B \log (r)$, since over a considerable range, loss is either spherical, cylindrical, or transitional. While more complex equations can readily be devised which will fit the data and include more of the variance, they add more complexity without increasing comparison testing effectiveness appreciably.

The analysis begins with the assumption that an intensity curve $X(r)$ of the type depicted in Fig. A1 can be represented in the form

$$X(r_k) = A + B \log (r_k) + X'(r_k), \tag{A1}$$

where $[r_k]$ $(k = 1, 2, \ldots N)$ is the sequence of range values. An application of standard methods, [A1], yields the following expressions for estimates of the coefficients, and the residual variance. Note the subscript, $e$, showing the estimate, as distinct from the true value, is shown only initially throughout the following material.

$$B_e = \frac{\sum_k \log (r_k) X(r_k) - [(\sum_k \log (r_k)) (\sum_k X(r_k)/N)]}{\sum_k (\log (r_k))^2 - [(\sum_k \log (r_k))^2/N]}, \tag{A2}$$

$$A_e = \frac{1}{N} (\sum_k X(r_k) - B_e \sum_k \log (r_k)), \tag{A3}$$

$$S_e^2 = \frac{\sum_k (X(r_k) - A_e - B_e \log (r_k))^2}{N - 2}. \tag{A4}$$

The type of test used to compare intensity curves in regard to long-term trend depends upon whether there is a significant oscillatory component in the residual curve $X'(r)$. If a test shows the residual values mutually independent, comparison tests based on the methods of linear regression analysis will apply. The tests are slightly weakened but still useful if significant oscillations are present. In special, demanding, cases the techniques of Eq. (A4) can be used to remove the oscillating component from the trend residual. The random residual remaining may be combined with the initial trend estimate and coefficients, Eqs. (A2), (A3), and (A4) redetermined. Before proceeding with the comparison tests, it is necessary to decide whether the trend residual is a random variable. A statistical test, a description of which follows, devised by M.G. Kendall [A2] is recommended for this purpose, because of its simplicity and effectiveness.

## A TEST FOR RANDOMNESS

The turning point test is based on the statistical hypothesis that the values $\{X'(r_k)\}$ $(k = 1, 2, \ldots n)$ are mutually independent; thus, they could have occurred in any order, each order being equally likely. An observed valued $x(r_k)$ is called a turning point if

$$X(r_{k-1}) < X(r_k) \text{ and } X(r_k) > X(r_{k+1}),$$

or if

$$X(r_{k-1}) > X(r_k) \text{ and } X(r_k) < X(r_{k+1}). \tag{A5}$$

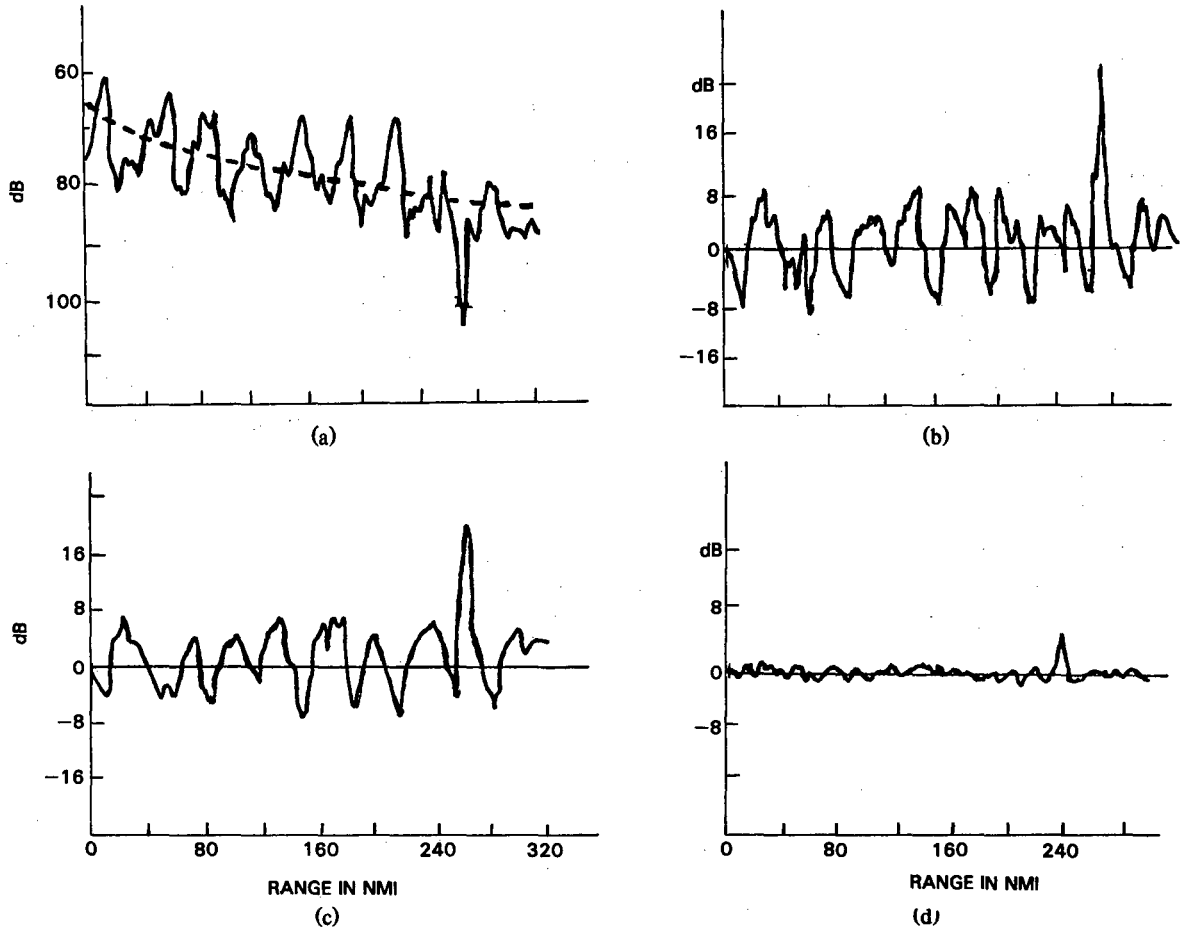



Fig. A1 — (a) A three-component propagation loss curve with an exponentially decaying average overlay (dashed line); (b) the derived oscillatory curve, $X'(r)$, after the trend in Fig. 1(a) is removed; Fig. (c) an autoregressive estimate, $X_0(r)$, of the residual oscillations in Fig. 1(b); Fig. (d) the final residual, $X_R(r)R$, equaling the difference between the oscillatory residual of Fig. 1(b) and the oscillatory estimate of Fig. 1(c).

Let $n_T$ denote the number of turning points which occur in a time series of $n$ distinct points. Assuming the above hypothesis, Kendall has shown [A2,p. 22-24] that for fairly large sample sizes, $n_T$ is approximately distributed as a normal random variable with mean

$$\mu = \frac{2}{3}(n - 2), \tag{A6}$$

and standard deviation

$$\sigma = \frac{16n - 29}{90}. \tag{A7}$$

The test procedure is: Select a confidence level $\alpha$. Reduce the series by removing repeated values, leaving $n$ distinct points, without changing their order of occurrence. Calculate $\mu$ and $\sigma$ by using Eqs. (A6) and (A7). Then the 100 $(1 - \alpha)$-percent confidence limits for $n_T$ are

$$\mu \pm \sigma Z_{\alpha/2}, \tag{A8}$$

where $Z\alpha/2$ denotes a percentage point of the normal distribution. Count the observed number $n_T$ of turning points for the series of distinct values. If $n_T$ is within the interval, we accept the hypothesis and conclude that the curve has no significant oscillatory component. Therefore, $X(r)$ consists only of

a long-term trend and a residual, random component. This residual series may not be a purely random process, but the oscillations it exhibits are not significant at the selected confidence level to warrant description. If $n_T$ lies outside the confidence interval, then we reject the hypotheses, and conclude that the series $X'(r)$ has a significant oscillatory component, which should be measured separately. The probability of making this decision when in fact the hypothesis is true is $\alpha$.

## TREND COMPARISON TESTS

Assume the turning point test has shown the residual curve $X'(r)$ to be random at same acceptable confidence level. By setting $z = \log(r)$ and $E(r) = X'(r)$, Eq. (A1) can be recast in the form

$$X = A + Bz + E, \tag{A9}$$

and we can apply the results of linear regression analysis (see for example, Section 22.9 of Ref. A1, Chapter 11 of Burr Ref. A3) to find confidence limits for $A$, $B$ and the standard error of estimate. At the 100 $(1 - \alpha)$ percent confidence level, we can calculate these limits as:

for $B$, the confidence limits are $B_e \pm L_B$ where,

$$L_B = \frac{t\left(\frac{\alpha}{2},\ N-2\right) S_e}{\left[\sum_k (\log(r_k) - (\sum_k \log(r_k)/N))^2\right]^{1/2}}; \tag{A10}$$

and

for $A$, the limits are $A_e \pm L_A$ where

$$L_A = t\left(\frac{\alpha}{2},\ N-2\right) S_e \left[\frac{\sum_k (\log(r_k))^2}{N \sum_k (\log(r_k) - (\sum_k \log(r_k)/N))^2}\right]. \tag{A11}$$

Here $t\ (\alpha/2,\ N-2)$ is a percentage point of the student $t$ distribution.

The standard deviation of the limits are

$$S_- = S_e \left[\frac{N-2}{X^2_{\alpha_2,\ N-2}}\right]^{1/2} \text{ and } S_+ = S_e \left[\frac{N-2}{X_{1-\frac{\alpha}{2},\ N-2}}\right]^{1/2}. \tag{A12}$$

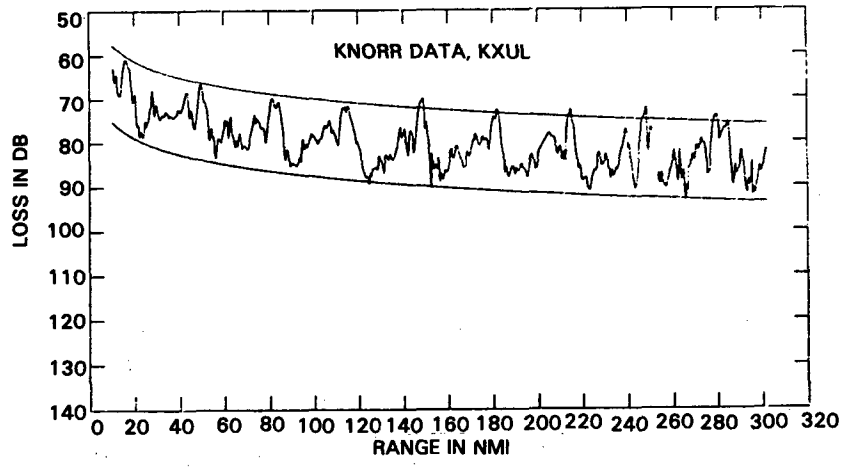
Here $X^2\ \mu,\ n$ is a percentage point of the Chi-square distribution.

In addition to the analytic comparisons previously described, a visual comparison plays the same qualitatively useful role as in traditional data/model comparisons. One such scheme used here is to superimpose the regression equation derived from one member of a comparison pair onto the data of the other member. To guide such visual comparisons, two displaced regression curves are used, separated by four residual standard deviations of the regression data. Equation (A13) shows the equation with subscripts, $E$, indicating experimental bounds, $L$, as illustrated in Figs. A2 and A3.

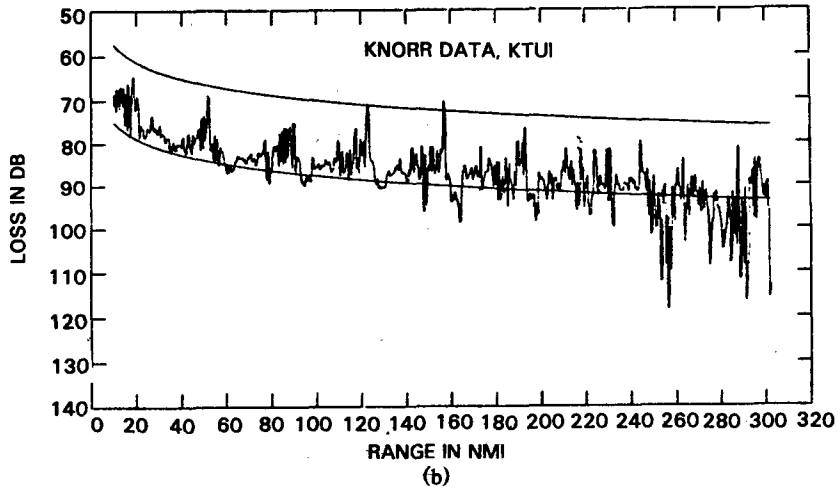
$$L = A_E + B_E \log(r) \pm 2\ S_E. \tag{A13}$$

With the plotted band shown on the figures we have computed an elementary overlap type measure called a Band-Fit (BF) coefficient as shown in Eq. (A14),

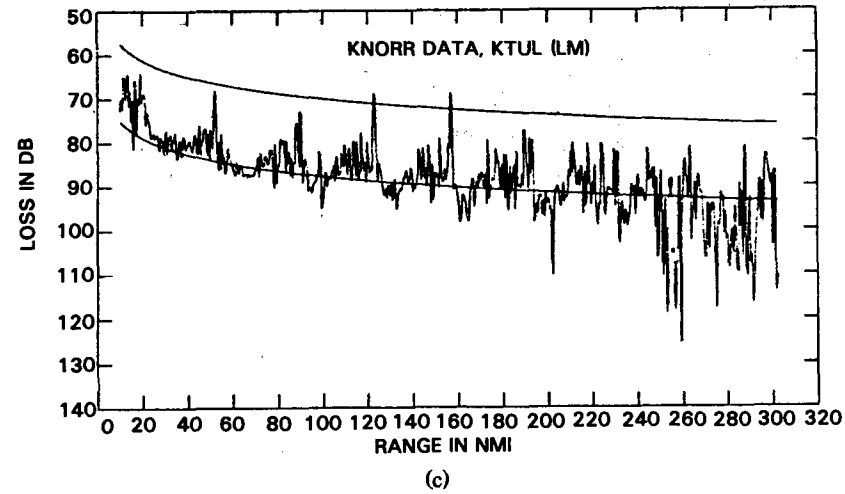
$$BF = \frac{P_M}{P_E}. \tag{A14}$$




(a)

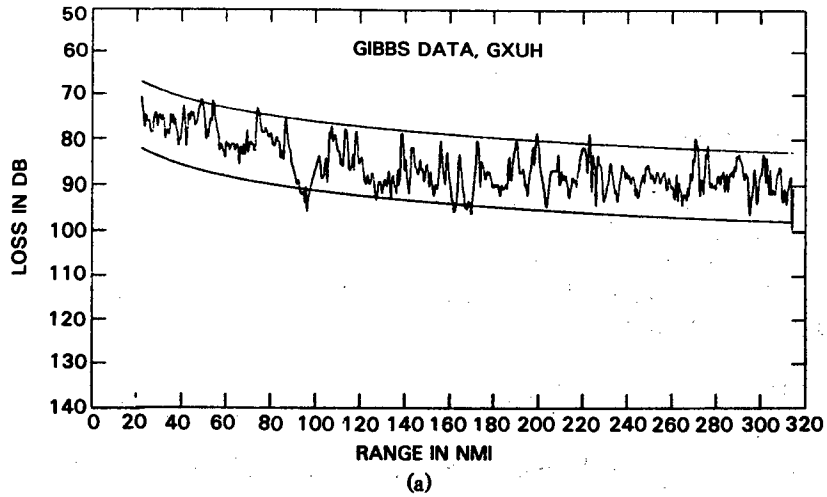



(b)

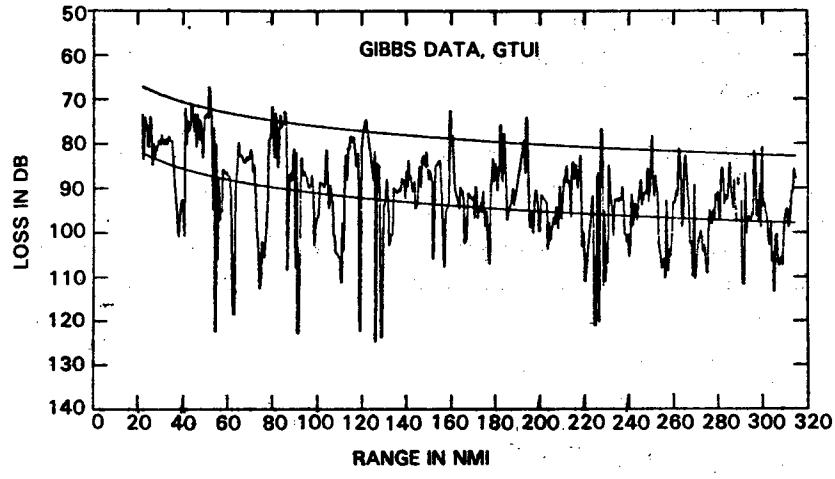



(c)

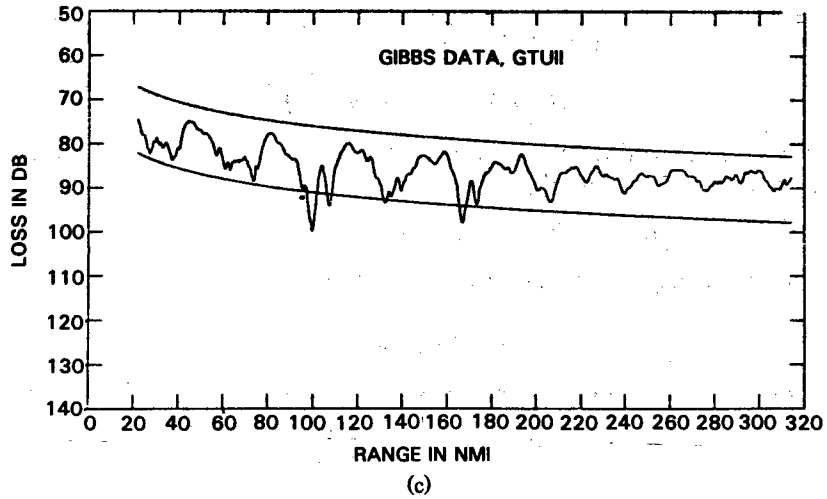
Fig. A2 — (a) The measured and fully corrected estimates of propagation loss taken on board the R/V KNORR; (b) TRIMAIN model calculations employing the bathymetric and velocity profile information acquired during the measurement of Fig. 2(a); (c) A second TRIMAIN calculation followed by the application of a Lloyd Mirror correction; (both model calculations are produced with deliberate flaws in the number of rays included beyond 250 km to provide material for the analysis technique to discriminate).




(a)




(b)




(c)

Fig. A3 — (a) The measured and fully corrected estimates of propagation loss taken on board the USNS GIBBS; (b) model calculations utilizing TRIMAIN and employing the same bathymetric and velocity profile information acquired during the measurement of Fig. 3(a); (c) A second TRIMAIN calculation followed by the application of a Lloyd Mirror correction.

The $P_M$ is the percentage of model points (as in later illustrations) which fall within the band superimposed and defined by the experimental data. The denominator, $P$, is the percentage of experimental data that is within the four-sigma band (Figs. A2 and A3).

## SEPARATION INTO OSCILLATORY AND RANDOM RESIDUAL COMPONENTS

After the residual curve, $X'(r)$, has been shown nonrandom by the turning point test, the oscillatory component must be separated from the final random residual. An autoregressive scheme was chosen to meet this need because of its effectiveness, and the suitability of the auto covariance function. Our discussion of autoregressive processes follows that of Ref. [A4], where complete derivations of the equations used can be found.

To begin with, an autoregressive process of order $m$, is defined as a second order uniformly sampled stationary random process $\{X(k)\}$ with zero mean, which satisfies the equation.

$$Y(k) = a_1\ Y(k-1) + a_2\ Y(k-2) + \ldots\ a_m\ Y(k-m) + Z(k) \tag{A15}$$

where $\{Z(k)\}$ is a purely random process. Here the coefficients $a_1, a_2, \ldots a_m$ are constant, and Eq. (A15) must hold for all observed values $k = 1, 2, \ldots N$. Note that an autoregressive process consists of two parts. The first, involving the coefficients $a_j$ is called the autoregressive scheme, and the second is called the residual process.

An autoregressive process may be generated by selecting an order $m$, a set of coefficients $\{a_j\}$ $(j = 1, 2, \ldots m)$ which satisfy a stationary condition, and a process $\{Z(k)\}$ obtained for example, from a table of independent normal deviates. Conversely, if one is given a process $\{Y(k)\}$, then one can attempt to fit an autoregressive process to $\{Y(k)\}$ in the following manner. Estimates $a_1, a_2, \ldots a_m$ of the autoregressive scheme coefficients are obtained as the solution of the system of $m$ equations

$$\begin{aligned} c_y(1) &= \hat{a}_1\ c_y(0) + \hat{a}_2\ c_y(-1) + \ldots\ a_m\ c_y(1-m), \\ c_y(2) &= \hat{a}_1\ c_y(1) + \hat{a}_2\ c_y(c) + \ldots\ \hat{a}_m\ c_y(2-m), \\ c_y(m) &= \hat{a}_1\ c_y(m-1) + \hat{a}_2\ c_y(m-2) + \ldots\ \hat{a}_m\ c_y(0), \end{aligned} \tag{A16}$$

where $c_y(k)$ is the sample autocovariance of the process $\{Y(k)\}$.

$$c_y(k) = \frac{1}{N} \sum_{j=1}^{N-k} Y(j)\ Y(j+k)$$

$$(k = 0, 1, \ldots, N-1). \tag{A17}$$

Note that $c_y(0)$ will give us an estimate of the variance of $\{Y(k)\}$. It may be shown that the variance of the residual process may be estimated by

$$S_z^2 = c_y(0) - \hat{a}_1\ c_y(1) - \ldots - \hat{a}_m\ c_y(m). \tag{A18}$$

To compare these two variances, we will use the normalized mean square error,

$$E_0 = S_z^2 / c_y(0). \tag{A19}$$

After the coefficients $\{a_1, a_2, \ldots a_m\}$ have been calculated, the residual process is obtained by subtracting the autoregressive scheme from $\{Y(k)\}$. To check whether a valid fit has been made, the residual should be tested to determine if it is purely random.

In selecting a time series model of this type, we are carrying out a program originally suggested in a paper by Whittle [A5]. He argues that any zero mean, stationary process whose spectral density satisfies a certain condition may be represented by an autoregression of infinite order. For such a process, a reasonably accurate estimate of the residual variance may be obtained by fitting a finite autore-

gressive scheme of sufficiently high order. The spectral condition requires that the reciprocal of the spectrum be expandable in a Fourier series (for all practical purposes, that the spectrum be nowhere zero), and is usually satisfied in practice.

Our analysis procedure then, is to calculate autoregressive scheme estimates for the different curves under consideration, using increasing values for $m$, calculating the normalized mean square error each time. Currently available computer codes (Robinson [A4] Section 2.8) enable us to do this with a minimum of time and effort. We can thus determine a value $m_0$ for $m$, such that the reduction in $E_0$ for higher order fits is insignificant in all cases.

For comparison, autoregressive fits of order $m_0$ are then used for all curves being analyzed. Granted that this requires an excessive number of terms in some cases, it provides a basis for comparison, without essentially affecting the estimate of the residual variance.

## MEASUREMENTS FOR OSCILLATORY COMPONENTS

Suppose now that the curve $X'(r)$ has been expressed as the sum of an oscillatory component $X_0(r)$ and a residual component $X_R(r)$. The oscillation can usually be attributed to some known physical cause such as the convergence zone effect. To study this phenomenon quantitatively, we next obtain a measure of this oscillatory component. For this, the sample autocovariance function defined by Eq. (A.6) is used. Thus, we calculate

$$c_y(k) = \frac{1}{N} \sum_{j=1}^{N-k} X_0\,(r_j)\; X_0\,(r_{j+k}), \tag{A20}$$

for $k = 0, 1, \ldots N-1$. A typical graph of $c_y(k)$ would resemble that of damped oscillatory motion starting at $k = 0$, with the variance $c_y(0)$ decreasing in magnitude as $k$ increases. In most cases the autocovariance function will be asymmetric or scalloped reflecting convergence zones, Lloyd mirror variation, or other origins most of which produce periodic but not sinusoidal variation. The principal period of the process is simply the distance between peaks of the function. To be specific, we will call this quantity the zone period, $P$. If we have oscillatory components for curves $C_1$, $C_2$ with zone periods $P_1$, $P_2$, then we may consider the zone spacing ratio, $Z$, defined by

$$Z = \frac{P_1 - P_2}{P_1}. \tag{A21}$$

A positive value of $z$ indicates that the $C_2$ oscillation has a shorter period than that of $C_1$.

## DISTRIBUTION OF VARIANCES

Let us briefly review the separation procedure which has been proposed for intensity curves: Starting with an initial curve $X$, a long-term trend $X_L$, is removed, leaving a residual curve $X'$. The residual curve is then decomposed as the sum of an oscillatory component $X_0$ and a random residual component $X_R$. We will denote by $V$, $V_L$, $V'$, $V_0$ and $V_R$ the variances of the previous curves. Because the three component series are uncorrelated, we will have

$$V' = V_0 + V_R, \tag{A22}$$

and

$$V = V_L + V_0 + V_R. \tag{A23}$$

Thus, the fractions $V_L/V$, $V_0/V$, and $V_R/V$ will adequately describe the distribution of the variance of the initial curve. One measure of the validity of the separation process is the extent to which Eqs. (A.22) and (A.23) hold. In all of the applications of the procedure examined values very close to theoretical predictions were observed. (See Table 8 of this report).

## REFERENCES

A1. M.G. Kendall, *The Advanced Theory of Statistics*, 4th ed., Vol. 2, Griffin, London, 1948.

A2. I.W. Burr, *Applied Statistical Methods,* Academic Press, New York, 1974.

A3. P. Whittle, "Tests of Fit in Time Series," *Biometrika* **39**, 309-319 (1952).

A4. E.A. Robinson, *Multichannel Time Series Analysis with Digital Computer Programs*, Rev. ed. Holden-Day, 1967.